(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 798 378 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
01.10.1997 Bulletin 1997/40

(51) Int Cl.6: C12N 15/12, C12N 15/62, C12N 15/85, C07K 14/72, C12N 1/21, C12N 5/16, C12Q 1/00, C12Q 1/68

(21) Application number: 97200903.9

(22) Date of filing: 25.03.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: 22.11.1996 EP 96203284
26.03.1996 EP 96200820

(71) Applicant: Akzo Nobel N.V.
6824 BM Arnhem (NL)

(72) Inventors:
- Mosselman, Sietse
5346 VM OSS (NL)
- Dijkema, Rein
5345 ML Oss (NL)

(74) Representative:
Ogilvie-Emanuelson, Claudia Maria et al
Patent Department Pharma
N.V. Organon
P.O. Box 20
5340 BH Oss (NL)

(54) Estrogen receptor

(57) The present invention relates to isolated DNA encoding novel estrogen receptors, the proteins encoded by said DNA, chimeric receptors comprising parts of said novel receptors and uses thereof.

EP 0 798 378 A2

Description

This invention relates to the field of receptors belonging to the superfamily of nuclear hormone receptors, in particular to steroid receptors. The invention relates to DNA encoding a novel steroid receptor, the preparation of said receptor, the receptor protein, and the uses thereof.

Steroid hormone receptors belong to a superfamily of nuclear hormone receptors involved in ligand-dependent transcriptional control of gene expression. In addition, this superfamily consists of receptors for non-steroid hormones such as vitamine D, thyroid hormones and retinoids (Giguère et al, Nature 330, 624-629, 1987; Evans, R.M., Science 240, 889-895,1988). Moreover, a range of nuclear receptor-like sequences have been identified which encode socalled 'orphan' receptors: these receptors are structurally related to and therefore classified as nuclear receptors, although no putative ligands have been identified yet (B.W. O'Malley, Endocrinology 125, 1119-1170, 1989; D.J. Mangelsdorf and R.M. Evans, Cell, 83, 841-850, 1995).

The superfamily of nuclear hormone receptors share a modular structure in which six distinct structural and functional domains, A to F, are displayed (Evans, Science 240, 889-895, 1988). A nuclear hormone receptor is characterized by a variabel N-terminal region (domain A/B), followed by a centrally located, highly conserved DNA-binding domain (hereinafter referred to as DBD; domain C), a variable hinge region (domain D), a conserved ligand-binding domain (herein after referred to as LBD; domain E) and a variable C-terminal region (domain F).

The N-terminal region, which is highly variable in size and sequence, is poorly conserved among the different members of the superfamily. This part of the receptor is involved in the modulation of transcription activation (Bocquel et al, Nucl. Acid Res., 17, 2581-2595, 1989; Tora et al, Cell 59, 477-487, 1989).

The DBD consists of approximately 66 to 70 amino acids and is responsible for DNA-binding activity: it targets the receptor to specific DNA sequences called hormone responsive elements (hereinafter referred to as HRE) within the transcription control unit of specific target genes on the chromatin (Martinez and Wahli, In 'Nuclear Hormone Receptors', Acad. Press, 125-153, 1991).

The LBD is located in the C-terminal part of the receptor and is primarily responsible for ligand binding activity. In this way, the LBD is essential for recognition and binding of the hormone ligand and, in addition possesses a transcription activation function, thereby determining the specificity and selectivity of the hormone response of the receptor. Although moderately conserved in structure, the LBD's are known to vary considerably in homology between the individual members of the nuclear hormone receptor superfamily (Evans, Science 240, 889-895, 1988; P.J. Fuller, FASEB J., 5, 3092-3099, 1991; Mangelsdorf et al. Cell, Vol. 83, 835-839, 1995).

Functions present in the N-terminal region, LBD and DBD operate independently from each other and it has been shown that these domains can be exchanged between nuclear receptors (Green et al, Nature, Vol. 325, 75-78, 1987). This results in chimeric nuclear receptors, such as described for instance in WO-A-8905355.

When a hormone ligand for a nuclear receptor enters the cell by diffusion and is recognized by the LBD, it will bind to the specific receptor protein, thereby initiating an allosteric alteration of the receptor protein. As a result of this alteration the ligand/receptor complex switches to a transcriptionally active state and as such is able to bind through the presence of the DBD with high affinity to the corresponding HRE on the chromatin DNA (Martinez and Wahli, 'Nuclear Hormone Receptors',125-153, Acad. Press, 1991). In this way the ligand/receptor complex modulates expression of the specific target genes. The diversity achieved by this family of receptors results from their ability to respond to different ligands.

The steroid hormone receptors are a distinct class of the nuclear receptor superfamily, characterized in that the ligands are steroid hormones. The receptors for glucocorticoids (GR), mineralcorticoids (MR), progestins (PR), androgens (AR) and estrogens (ER) are classical steroid receptors. Furthermore, the steroid receptors have the unique ability upon activation to bind to palindromic DNA sequences, the so-called HRE's, as homodimers. The GR, MR, PR and AR recognize the same DNA sequence, while the ER recognizes a different DNA sequence. (Beato et al, Cell, Vol. 83, 851-857, 1995). After binding to DNA, the steroid receptor is thought to interact with components of the basal transcriptional machinery and with sequence-specific transcription factors, thus modulating the expression of specific target genes.

Several HRE's have been identified, which are responsive to the hormone/receptor complex. These HRE's are situated in the transcriptional control units of the various target genes such as mammalian growth hormone genes (responsive to glucocorticoid, estrogen, testosterone), mammalian prolactin genes and progesterone receptor genes (responsive to Estrogen), avian ovalbumin genes (responsive to progesterone), mammalian metallothionein gene (responsive to glucocorticoid) and mammalian hepatic $\alpha_{2\mu}$-globulin gene (responsive to estrogen, testosterone, glucocorticoid).

The steroid hormone receptors have been known to be involved in embryonic development, adult homeostasis as well as organ physiology. Various diseases and abnormalities have been ascribed to a disturbance in the steroid hormone pathway. Since the steroid receptors exercise their influence as hormone-activated transcriptional modulators, it can be anticipated that mutations and defects in these receptors, as well as overstimulation or blocking of these

receptors might be the underlying reason for the altered pattern. A better knowledge of these receptors, their mechanism of action and of the ligands which bind to said receptor might help to create a better insight in the underlying mechanism of the hormone signal transduction pathway, which eventually will lead to better treatment of the diseases and abnormalities linked to altered hormone/receptor functioning.

For this reason cDNA's of the steroid and several other nuclear receptors of several mammalians, including humans, have been isolated and the corresponding amino acid sequences have been deduced, such as for example the human steroid receptors PR, ER, GR, MR and AR, the human non-steroid receptors for vitamine D, thyroid hormones, and retinoids such as retinol A and retinoic acid. In addition, cDNA's encoding well over 100 mammalian orphan receptors have been isolated, for which no putative ligands are known yet (Mangelsdorf et al, Cell, Vol.83, 835-839, 1995). However, there is still a great need for the elucidation of other nuclear receptors in order to unravel the various roles these receptors play in normal physiology and pathology.

The present invention provides for such a novel nuclear receptor. More specific, the present invention provides for novel steroid receptors, having estrogen mediated activity. Said novel steroid receptors are novel estrogen receptors, which are able to bind and be activated by, for example, estradiol, estrone and estriol.

According to the present invention it has been found that a novel estrogen receptor is expressed as an 8 kb transcript in human thymus, spleen, peripheral blood lymphocytes (PBLs), ovary and testis. Furthermore, additional transcripts have been identified. Another transcript of approximately 10 kb was identified in ovary, thymus and spleen. In testis, an additional transcript of 1.3 kb was detected. These transcripts are probably generated by alternative splicing of the gene encoding the novel estrogen receptor according to the invention.

Cloning of the cDNA's encoding the novel estrogen receptors according to the invention revealed that several splicing variants of said receptor can be distinguished. At the protein level, these variants differ only at the C-terminal part.

cDNA encoding an ER has been isolated (Green, et al, Nature 320, 134-139, 1986; Greene et al, Science 231, 1150-1154, 1986), and the corresponding amino acid sequence has been deduced. This receptor and the receptor according to the present invention, however, are distinct, and encoded for by different genes with different nucleic acid sequences. Not only do the ER of the prior art (hereinafter referred to as classical ER) and the ER according to the present invention differ in amino acid sequence, they also are located on different chromosomes. The gene encoding the classical ER is located on chromosome 6, whereas the gene encoding the ER according to the invention was found to be located on chromosome 14. The ER according to the invention furthermore distinguishes itself from the classical receptor in differences in tissue distribution, indicating that there may be important differences between these receptors at the level of estrogenic signalling.

In addition, two orphan receptors, ERRα and ERRβ, having an estrogen receptor related structure have been described (Giguère et al, Nature 331, 91-94, 1988). These orphan receptors, however, have not been reported to be able to bind estrodial or any other hormone that binds to the classical ER, and other ligands which bind to these receptors have not been found yet. The novel estrogen receptor according to the invention distinguishes itself clearly from these receptors since it was found to bind estrogens.

The fact that a novel ER according to the invention has been found is all the more surprising, since any suggestion towards the existence of additional estrogen receptors was absent in the scientific literature: neither the isolation of the classical ER nor the orphan receptors ERRα and ERRβ suggested or hinted towards the presence of additional estrogen receptors such as the receptors according to the invention. The identification of additional ER's could be a major step forward for the existing clinical therapies, which are based on the existence of one ER and as such ascribe all estrogen mediated abnormalities and/or diseases to this one receptor. The receptors according to the invention will be useful in the development of hormone analogs that selectively activate either the classical ER or the novel estrogen receptor according to the invention. This should be considered as one of the major advantages of the present invention.

Thus, in one aspect, the present invention provides for isolated cDNA encoding a novel steroid receptor. In particular, the present invention provides for isolated cDNA encoding a novel estrogen receptor.

According to this aspect of the present invention, there is provided an isolated DNA encoding a steroid receptor protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said receptor protein exhibits at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and the amino acid sequence of said ligand-binding domain of said receptor protein exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4.

In particular, the isolated DNA encodes a steroid receptor protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said receptor protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO:3.

More particularly, the isolated DNA encodes a steroid receptor protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain , wherein the amino acid sequence of said ligand-binding domain of said receptor protein exhibits at least 75%, preferably 80%, more preferably 90%, most preferably 100% homology with the amino

acid sequence shown in SEQ ID NO:4.

A preferred isolated DNA according to the invention encodes a steroid receptor protein having the amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.

A more preferred isolated DNA according to the invention is an isolated DNA comprising a nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20 or SEQ ID NO:24.

The DNA according to the invention may be obtained from cDNA. Alternatively, the coding sequence might be genomic DNA or prepared using DNA synthesis techniques.

The DNA according to the invention will be very useful for in vivo expression of the novel receptor proteins according to the invention in sufficient quantities and in substantially pure form.

In another aspect of the invention there is provided for a steroid receptor comprising the amino acid sequence encoded by the above described DNA molecules.

The steroid receptor according to the invention has an N-terminal domain, a DNA-binding domain and a ligand-binding domain wherein the amino acid sequence of said DNA-binding domain of said receptor exhibits at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and the amino acid sequence of said ligand-binding domain of said receptor exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4.

In particular the steroid receptor according to the invention has an N-terminal domain, a DNA-binding domain and a ligand-binding domain wherein the amino acid sequence of said DNA-binding domain of said receptor exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO 3

More particular, the steroid receptor according to the invention has an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said ligand-binding domain of said receptor exhibits at least 75%, prefearbly 80%, more preferably 90%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO 4

It will be clear for those skilled in the art that also steroid receptor proteins comprising combined DBD and LBD preferences and DNA encoding such receptors are subject of the invention.

Preferably, the steroid receptor according to the invention comprises an amino acid sequence shown in SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.

Also within the scope of the present invention are steroid receptor proteins which comprise variations in the amino acid sequence of the DBD and LBD without loosing their respective DNA-binding or ligand-binding activities. The variations that can occur in those amino acid sequences comprise deletions, substitutions, insertions, inversions or additions of (an) amino acid(s) in said sequence, said variations resulting in amino acid difference(s) in the overall sequence. It is well known in the art of proteins and peptides that these amino acid differences lead to amino acid sequences that are different from, but still homologous with the native amino acid sequence they have been derived from.

Amino acid substitutions that are expected not to essentially alter biological and immunological activities, have been described in for example Dayhof, M.D., Atlas of protein sequence and structure, Nat. Biomed. Res. Found., Washington D.C., 1978, vol. 5, suppl. 3. Amino acid replacements between related amino acids or replacements which have occurred frequently in evolution are, inter alia Ser/Ala, Ser/Gly, Asp/Gly, Arg/Lys, Asp/Asn, Ile/Val. Based on this information Lipman and Pearson developed a method for rapid and sensitive protein comparison (Science 227, 1435-1441, 1985) and determining the functional similarity between homologous polypeptides.

Variations in amino acid sequence of the DBD according to the invention resulting in an amino acid sequence that has at least 80% homology with the sequnece of SEQ ID NO:3 will lead to receptors still having sufficient DNA binding activity. Variations in amino acid sequence of the LBD according to the invention resulting in an amino acid sequence that has at least 70% homology with the sequnece of SEQ ID NO:4 will lead to receptors still having sufficient ligand binding activity.

Homology as defined herein is expressed in percentages, determined via PCGENE. Homology is calculated as the percentage of identical residues in an alignment with the sequence according to the invention. Gaps are allowed to obtain maximum alignment.

Comparing the amino acid sequences of the classical ER and the ER's according to the invention revealed a high degree of similarity within their respective DBD's. The conservation of the P-box (amino acids E-G-X-X-A) which is responsible for the actual interactions of the classical ER with the target DNA element (Zilliacus et al., Mol.Endo. 9, 389, 1995; Glass. End.Rev. 15, 391, 1994), is indicative for a recognition of estrogen responsive elements (ERE's) by the ER's according to the invention. The receptors according to the invention indeed showed ligand-dependent trans-activation on ERE-containing reporter constructs. Therefore, the classical ER and the novel ER's according to the invention may have overlapping target gene specificities. This could indicate that in tissues which co-express both respective ER's, these receptors compete for ERE's. The ER's according to the invention may regulate transcription of target genes differently from classical ER regulation or could simply block classical ER functioning by occupying estrogen responsive elements. Alternatively, transcription might be influenced by heterodimerization of the different

receptors.

Thus, a preferred steroid receptor according to the invention comprises the amino acid sequence E-G-X-X-A within the P box of the DNA binding domain, wherein X stands for any amino acid. Also within the scope of the invention is isolated DNA encoding such a receptor.

Methods to prepare the receptors according to the invention are well known in the art (Sambrook et al., Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory Press. Cold Spring Harbor, 1989). The most practical approach is to produce these receptors by expression of the DNA encoding the desired protein.

A wide variety of host cell and cloning vehicle combinations may be usefully employed in cloning the nucleic acid sequence coding for the receptor of the invention. For example, useful cloning vehicles may include chromosomal, non-chromosomal and synthetic DNA sequences such as various known bacterial plasmids and wider host range plasmids and vectors derived from combinations of plasmids and phage or virus DNA. Useful hosts may include bacterial hosts, yeasts and other fungi, plant or animal hosts, such as Chinese Hamster Ovary (CHO) cells or monkey cells and other hosts.

Vehicles for use in expression of the ligand-binding domain of the present invention will further comprise control sequences operably linked to the nucleic acid sequence coding for the ligand-binding domain. Such control sequences generally comprise a promoter sequence and sequences which regulate and/or enhance expression levels. Furthermore an origin of replication and/or a dominant selection marker are often present in such vehicles. Of course control and other sequences can vary depending on the host cell selected.

Techniques for transforming or transfecting host cells are quite known in the art (see, for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, 1989).

Recombinant expression vectors comprising the DNA of the invention as well as cells transformed with said DNA or said expression vector also form part of the present invention.

In a further aspect of the invention, there is provided for a chimeric receptor protein having an N-terminal domain, a DNA-binding domain, and a ligand-binding domain, characterized in that at least one of the domains originates from a receptor protein according to the invention, and at least one of the other domains of said chimeric protein originates from another receptor protein from the nuclear receptor superfamily, provided that the DNA-binding domain and the ligand-binding domain of said chimeric receptor protein originate from different proteins.

In particular, the chimeric receptor according to the invention comprises the LBD according to the invention, said LBD having an amino acid sequence which exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4. In that case the N-terminal domain and DBD should be derived from another nuclear receptor, such as for example PR. In this way a chimeric receptor is constructed which is activated by a ligand of the ER according to the invention and which targets a gene under control of a progesterone responsive element. The chimeric receptors having a LBD according to the invention are useful for the screening of compounds to identify novel ligands or hormone analogs which are able to activate an ER according to the invention.

In addition, chimeric receptors comprising a DBD according to the invention, said DBD having an amino acid sequence exhibiting at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and a LBD and, optionally, an N-terminal domain derived from another nuclear receptor, can be succesfully used to identify novel ligands or hormone analogs for said nuclear receptors. Such chimeric receptors are especially useful for the identification of the respective ligands of orphan receptors.

Since steroid receptors have three domains with different functions, which are more or less independent, it is possible that all three functional domains have been derived from different members of the steroid receptor superfamily.

Molecules which contain parts having a different origin are called chimeric. Such a chimeric receptor comprising the ligand-binding domain and/or the DNA-binding domain of the invention may be produced by chemical linkage, but most preferably the coupling is accomplished at the DNA level with standard molecular biological methods by fusing the nucleic acid sequences encoding the necessary steroid receptor domains. Hence, DNA encoding the chimeric receptor proteins according to the invention are also subject of the present invention.

Such chimeric proteins can be prepared by transfecting DHA encoding these chimeric receptor proteins to suitable host cells and culturing these cells under suitable conditions.

It is extremely practical if, next to the information for the expression of the steroid receptor, also the host cell is transformed or transfected with a vector which carries the information for a reporter molecule. Such a vector coding for a reporter molecule is characterized by having a promoter sequence containing one or more hormone responsive elements (HRE) functionally linked to an operative reporter gene. Such a HRE is the DNA target of the activated steroid receptor and, as a consequence, it enhances the transcription of the DNA, coding for the reporter molecule. In *in vivo* settings of steroid receptors the reporter molecule comprises the cellular response to the stimulation of the ligand. However, it is possible *in vitro* to combine the ligand-binding domain of a receptor to the DNA binding domain and transcription activating domain of other steroid receptors, thereby enabling the use of other HRE and reporter molecule systems. One such a system is established by a HRE presented in the MMTV-LTR (mouse mammary tumor virus long terminal repeat sequence in connection with a reporter molecule like the firefly luciferase gene or the bacterial gene

for CAT (chloramphenicol transferase). Other HRE's which can be used are the rat oxytocin promotor, the retinoic acid responsive element, the thyroid hormone responsive element, the estrogen responsive element and also synthetic responsive elements have been described (for instance in Fuller, ibid. page 3096). As reporter molecules next to CAT and luciferase β-galactosidase can be used.

Steroid hormone receptors and chimeric receptors according to the present invention can be used for the *in vitro* identification of novel ligands or hormonal analogs. For this purpose binding studies can be performed with cells transformed with DNA according to the invention or an expression vector comprising DNA according to the invention, said cells expressing the steroid receptors or chimeric receptors according to the invention.

The novel steroid hormone receptor and chimeric receptors according to the invention as well as the ligand-binding domain of the invention, can be used in an assay for the identification of functional ligands or hormone analogs for the nuclear receptors.

Thus, the present invention provides for a method for identifying functional ligands for the steroid receptors and chimeric receptors according to the invention, said method comprising the steps of

a) introducing into a suitable host cell 1) DNA or an expression vector according to the invention, and 2) a suitable reporter gene functionally linked to an operative hormone response element, said HRE being able to be activated by the DNA-binding domain of the receptor protein encoded by said DNA;
b) bringing the host cell from step a) into contact with potential ligands which will possibly bind to the ligand-binding domain of the receptor protein encoded by said DNA from step a);
c) monitoring the expression of the receptor protein encoded by said reporter gene of step a).

If expression of the reporter gene is induced with respect to basic expression (without ligand), the functional ligand can be considered as an agonist; if expression of the reporter gene remains unchanged or is reduced with respect to basic expression, the functional ligand can be a suitable (partial) antagonist.

For performing such kind of investigations host cells which have been transformed or transfected with both a vector encoding a functional steroid receptor and a vector having the information for a hormone responsive element and a connected reporter molecule are cultured in a suitable medium. After addition of a suitable ligand, which will activate the receptor the production of the reporter molecule will be enhanced, which production simply can be determined by assays having a sensitivity for the reporter molecule. See for instance WO-A-8803168. Assays with known steroid receptors have been described (for instance S. Tsai et al., Cell 57, 443, 1989; M. Meyer et al., Cell 57, 433, 1989).

Legends to the figures

Figure 1.

Northern analysis of the novel estrogen receptor (ERβ). Two different multiple tissue Northern blots (Clontech) were hybridised with a specific probe for ERβ (see examples). Indicated are the human tissues the RNA originated from and the position of the size markers in kilobases (kb).

Figure 2.

Histogram showing the 3- to 4-fold stimulatory effect of 17β-estradiol, estriol and estrone on the luciferase activity mediated by Enβ. An expression vector encoding ERβ was transiently transfected into CHO cells together with a reporter construct containing the rat oxytocin promoter in front of the firefly luciferase encoding sequence (see examples).

Figure 3.

Effect of 17β-estradiol (E2) alone or in combination with the anti-estrogen ICI-164384 (ICI) on ERα and ERβ. Expression constructs for ERα (the classical ER) and ERβ were transiently transfected into CHO cells together with the rat oxytocin promoter-luciferase reporter construct described in the examples. Luciferase activities were determined in triplicate and normalised for transfection efficiency by measuring β-galactosidase in the same lysate.

Figure 4.

Expression of ERα and ERβ in a number of cell lines determined by RT-PCR analysis (see examples). The cell lines used were derived from different tissues/cell types: endometrium (ECC1, Ishikawa, HEC-1A, RL95-2); osteosarcoma (SAOS-2, U2-OS, HOS, MG63); breast tumours (MCF-7, T47D), endothelium (HUV-EC-C, BAEC-1); smooth muscle (HISM, PAC-1, A7R5, A10, RASMC, CavaSMC); liver (HepG2); colon (CaCo2); and vagina (Hs-760T, SW-954).

All cell lines were human except for PAC-1, A7R5, A10 and RASMC which are of rat origin, BAEC-1 which is of bovine origin and CavaSMC which is of guinea pig origin

Figure 5.

Transactivation assay using stably transfected CHO cell lines expressing ERα or ERβ together with the rat oxytocin-luciferase estrogen-responsive reporter (see examples for details). Hormone-dependent transactivation curves were determined for 17β-estradiol and for Org4094. For the ER antagonist raloxifen, cells were treated with 2 x 10^{-10} mol/L 17β-estradiol together with increasing concentrations of raloxifen. Maximal values of the responses were arbitrarily set at 100%.

Examples

A. Molecular cloning of the novel estrogen receptor.

Two degenerate oligonucleotides containing inosines (I) were based on conserved regions of the DNA-binding domains and the ligand-binding domains of the human steroid hormone receptors.

Primer #1:

5'-GGIGA(C/T)GA(A/G)GC(A/T)TCIGGITG(C/T)CA(C/T)TA(C/T)GG-3' (SEQ ID NO:7).

Primer #2:

5'-AAGCCTGG(C/G)A(C/T)IC(G/T)(C/T)TTIGCCCAI(C/T)TIAT-3' SEQ ID NO:8).

As template, cDNA from human EBV-stimulated PBLs (peripheral blood leukocytes) was used. One microgram of total RNA was reverse transcribed in a 20 µl reaction containing 50 mM KCl, 10 mM Tris-HCl pH 8.3, 4 mM MgCl2, 1 mM dNTPs (Pharmacia), 100 pmol random hexanucleotides (Pharmacia), 30 Units RNAse inhibitor (Pharmacia) and 200 Units M-MLV Reverse transcriptase (Gibco BRL). Reaction mixtures were incubated at 37°C for 30 minutes and heat-inactivated at 100°C for 5 minutes. The cDNA obtained was used in a 100 µl PCR reaction containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl2, 0.001% gelatin (w/v), 3% DMSO, 1 microgram of primer #1 and primer #2 and 2.5 Units of Amplitaq DNA polymerase (Perkin Elmer). PCR reactions were performed in the Perkin Elmer 9600 thermal cycler. The initial denaturation (4 minutes at 94°C) was followed by 35 cycles with the following conditions: 30 sec. 94°C, 30 sec. 45°C, 1 minute 72°C and after 7 minutes at 72°C the reactions were stored at 4°C. Aliquots of these reactions were analysed on a 1.5% agarose gel. Fragments of interest were cut out of the gel, reamplified using identical PCR-conditions and purified using Qiaex II (Qiagen). Fragments were cloned in the pCRII vector and transformed into bacteria using the TA-cloning kit (Invitrogen). Plasmid DNA was isolated for nucleotide sequence analysis using the Qiagen, plasmid midi protocol (Qiagen). Nucleotide sequence analysis was performed with the ALF automatic sequencer (Pharmacia) using a T7 DNA sequencing kit (Pharmacia) with vector-specific or fragment-specific primers.

One cloned fragment corresponded to a novel estrogen receptor (ER) which is closely related to the classical estrogen receptor. Part of the cloned novel estrogen receptor fragment (nucleotides 466 to 797 in SEQ ID 1) was amplified by PCR using oligonucleotide #3 TGTTACGAAGTGGGAATGGTGA (SEQ ID NO:9) and oligonucleotide #2 and used as a probe to screen a human testis cDNA library in λgt11 (Clontech #HL1010b). Recombinant phages were plated (using Y1090 bacteria grown in LB medium supplemented with 0.2% maltose) at a density of 40.000 pfu (plaque-forming units) per 135 mm dish and replica filters (Hybond-N, Amersham) were made as described by the supplier. Filters were prehybridised in a solution containing 0.5 M phosphate buffer (pH 7.5) and 7% SDS at 65°C for at least 30 minutes. DNA probes were purified with Qiaex II (Qiagen), ^{32}P-labeled with a Decaprime kit (Ambion) and added to the prehybridisation solution. Filters were hybridised at 65°C overnight and then washed in 0.5 X SSC/0,1% SDS at 65°C. Two positive plaques were identified and could be shown to be identical. These clones were purified by rescreening one more time. A PCR reaction on the phage eluates with the λgt11-specific primers #4: 5'-TTGACACCAGACCAACTGGTAATG-3' (SEQ ID NO:10) and #5: 5'-GGTGGCGACGACTCCTGGAGCCCG-3' (SEQ ID NO:11)

yielded a fragment of 1700 basepairs on both clones.

Subsequent PCR reactions using combinations of a gene-specific primer #6: 5'-GTACACTGATTTGTAGCTGGAC-3' (SEQ ID NO:12) with the λgtll primer #4 and gene-specific primer #7: 5'-CCATGATGATGTCCCTGACC-3' (SEQ ID NO:13) with λgt11 primer primer #5 yielded fragments of approximately 450 bp and 1000 bp. respectively. which were cloned in the pCRII vector and used for nucleotide sequence analysis. The conditions for these PCR reactions were as described above except for the primer concentrations (200 ng of each primer) and the annealing temperature (60°C). Since in the cDNA clone the homology with the ER is lost abruptly at a site which corresponds to the exon 7/exon 8 boundary in the ER (between nucleotides 1247 and 1248 in SEQ ID NO:1), it was suggested that this sequence corresponds to intron 7 of the novel ER gene. For verification of the nucleotide sequences of this cDNA clone, a 1200 bp fragment was generated on the cDNA clone with λgt11 primer #4 with a gene-specific primer #8 corresponding to the 3' end of exon 7: 5'-TCGCATGCCTGACGTGGGAC-3' (SEQ ID NO:14) using the proofreading *Pfu* polymerase (Stratagene). This fragment was also cloned in the pCRII vector and completely sequenced and was shown to be identical to the sequences obtained earlier.

To obtain nucleotide sequences of the novel ER downstream of exon 7, a degenerate oligonucleotide based on the AF-2 region of the classical ER (#9: 5'-GGC(C/G)TCCAGCATCTCCAG(C/G)A(A/G)GAG-3'; SEQ ID NO:15) was used together with the gene-specific oligonucleotide #10: 5'-GGAAGCTGGCTCACTTGCTG-3' (SEQ ID NO:16) using testis cDNA as template (Marathon ready testis cDNA, Clontech Cat #7414-1). A specific 220 bp fragment corresponding to nucleotides 1112 to 1332 in SEQ ID No. 1 was cloned and sequenced. Nucleotides 1112 to 1247 were identical to the corresponding sequence of the cDNA clone. The sequence downstream thereof is highly homologous with the corresponding region in the classical ER. In order to obtain sequences of the novel ER downstream of the AF-2 region, RACE (rapid amplification of cDNA ends) PCR reactions were performed using the Marathon-ready testis cDNA (Clontech) as template. The initial PCR was performed using oligonucleotide #11: 5'-TCTTGTTCTGGACAGGGATG-3' (SEQ ID NO:17) in combination with the AP1 primer provided in the kit. A nested PCR was performed on an aliquot of this reaction using oligonucleotide #10 (SEQ ID NO:16) in combination with the oligo dT primer provided in the kit. Subsequently, an aliquot of this reaction was used in a nested PCR using oligonucleotide #12: 5'-GCATGGAACATCTGCTCAAC-3' (SEQ ID NO:18) in combination with the oligo dT primer. Nucleotide sequence analysis of a specific fragment that was obtained (corresponding to nucleotides 1256 to 1431 in SEQ ID NO 1) revealed a sequence encoding the carboxyterminus of the novel ER ligand-binding domain, including an F-domain and a translational stop codon and part of the 3' untranslated sequence which is not included in SEQ ID NO:1. The deduced amino acid sequence is shown in SEQ ID NO:5.

In order to investigate the possibility that the novel estrogen receptor had additional, upstream translation-initiation codons, RACE-PCR experiments were performed using Marathon-ready testis cDNA (Clontech Cat. # 7414-1). First a PCR was performed using oligonucleotide SEQ ID NO:12 (antisense corresponding to nucleotides 416-395 in SEQ ID NO:1) and AP-1 (provided in the kit). A nested PCR was then performed using oligonucleotide having SEQ ID NO: 27 (antisense corresponding to nucleotides 254-231 in SEQ ID NO:1) with AP-2 (provided in the kit). From the smear that was obtained, the region corresponding to fragments larger than 300 basepairs was cut out, purified using the GenecleanII kit (Bio101) and cloned using the TA-cloning kit (Clontech). Colonies were screened by PCR using gene-specific primers: SEQ ID NO:22 and SEQ ID NO:28. The clone containing the largest insert was sequenced. The nucleotide sequence corresponds to nucleotides 1 to 490 in SEQ ID NO:24. It is clear from this sequence that the first in-frame upstream translation initiation codon is present at position 77-79 in SEQ ID NO:24. Upstream of this translational startcodon an in-frame stop-codon is present (11-13 in SEQ ID NO:24). Consequently, the reading frame of the novel estrogen receptor is 530 amino acids (shown in SEQ ID NO:25) and has a calculated molecular mass of 59.234 kD.

To confirm the nucleotide sequences obtained by 5' RACE, human genomic clones were obtained and analysed. A human genomic library in λEMBL3 (Clontech HL1067J) was screened with a probe corresponding to nucleotides 1 to 416 in SEQ ID NO:1. A strongly hybridizing clone was plaque-purified and DNA was isolated using standard protocols (Sambrook et al, 1989). The DNA was digested with several restriction enzymes, electrophoresed on agarose gel and blotted onto Nylon filters. Hybridisation of the blot with a probe corresponding to the above-mentioned RACE fragment (nucleotides 1-490 in SEQ ID NO:24) revealed a hybridizing Sau3A fragment of approximately 800 basepairs. This fragment was cloned into the BamH1 site of pGEM3Z and sequenced. The nucleotide sequence contained one base difference which is probably a PCR-induced point mutation in the RACE fragment. Nucleotide 172 was a G residue in the 5'RACE fragment, but an A residue in several independent genomic subclones.

B. Identification of two splice variants of the novel estrogen receptor.

Rescreening of the testis cDNA library with a probe corresponding to nucleotides 918 to 1246 in SEQ ID No. 1 yielded two hybridizing clones, the 3' end of which were amplified by PCR (gene-specific primer #10: 5'-GGAAGCTGGCTCACTTGCTG-3' (SEQ ID NO:16) together with primer #4, SEQ ID NO:10), cloned and sequenced. One clone

was shown to contain an alternative exon 8 (exon 8B) of the novel ER. In SEQ ID No. 2 the protein encoding part and the stopcodon of this splice variant are presented As a consequence of the introduction of this exon through an alternative splicing reaction, the reading frame encoding the novel ER is immediately terminated, thereby creating a truncation of the carboxyterminus of the novel ER (SEQ ID NO:6).

Screening of a human thymus cDNA library (Clontech HL1074a) with the probe corresponding to nucleotides 918 to 1246 in SEQ ID No. 1, revealed another splice variant. The 3' end of one hybridizing clone was amplified using primer #10 (SEQ ID NO:16) with the λgt10-specific primer #13 5'-AGCAAGTTCAGCCTGTTAAGT-3' (SEQ ID NO:19), cloned and sequenced. The obtained nucleotide sequence upstream of the exon 7/exon 8 boundary was identical to the clones identified earlier. However, an alternative exon 8 (exon 8C) was present at the 3' end encoding two C-terminal amino acids followed by a stop-codon. The nucleotide sequence of the protein-encoding part of this splice variant is shown in SEQ ID NO:20, the corresponding protein sequence is SEQ ID NO:21.

These two variants of the novel estrogen receptor do not contain the AF-2 region and therefore probably lack the ability to modulate transcription of target genes in a ligand-dependent fashion. However, the variants potentially could interfere with the functioning of the wild-type classical ER and/or the wild-type novel ER, either by heterodimerization or by occupying estrogen response elements or by interactions with other transcription factors. A mutant of the classical ER (ER1-530) has been described which closely resembles the two variants of the novel estrogen receptor described above. ER1-530 has been shown to behave as a dominant-negative receptor i.e. it can modulate the intracellular activity of the wild type ER (Ince *et al*, J. Biol. Chem. 268, 14026-14032, 1993).

C. Northern blot analysis.

Human multiple tissue Northern blots (MTN-blots) were purchased from Clontech and prehybridized for at least 1 hour at 65°C in 0.5 M phosphate buffer pH 7.5 with 7% SDS. The DNA fragment that was used as a probe (corresponding to nucleotides 466 to 797 in SEQ ID No. 1) was ^{32}P-labeled using a labelling kit (Ambion), denatured by boiling and added to the prehybridisation solution. Washing conditions were: 3X SSC at room temperature, followed by 3 X SSC at 65°C, and finally 1 X SSC at 65°C. The filters were than exposed to X-ray films for one week. Two transcripts of approximately 8 kb and 10 kb were detected in thymus, spleen, ovary and testis. In addition, a 1.3 kb transcript was detected in testis.

D. RT-PCR analysis of expression of ERα and ERβ in cell lines.

RNA was isolated from a number of human and animal cell lines using RNAzol B (Cinna/Biotecx). cDNA was made using 2.5 microgram of total RNA using the Superscript II kit (BRL) following the manufacturers instructions. A portion of the cDNA was used for specific PCR amplifications of fragments corresponding either to mRNA encoding the ER or to the novel estrogen receptor. (It should be emphasized that the primers used are based on human and rat sequences, whereas some of the cell lines were not rat or human, see legend of Figure 4). Primers used were for ERα: sense 5'-GATGGGCTTACTGACCAACC-3' and antisense 5'-AGATGCTCCATGCCTTTG-3' generating a 548 base pair fragment corresponding to part of the LBD. For ERβ: sense 5'- TTCACCGAGGCCTCCATGATG-3' and antisense 5'-CAGATGTTCCATGCCCTTGTT-3' generating a 565 base pair fragment corresponding to part of the LBD. The PCR samples were analysed on agarose which were blotted onto Nylon membranes. These blots were hybridised with ^{32}P-labeled PCR fragments generated with the above-mentioned primers on ERα and ERβ plasmid DNA using standard experimental procedures (Sambrook *et al*, 1989).

E. Ligand-dependent transcription activation by the novel estrogen receptor protein.

Cell culture

Chinese Hamster Ovary (CHO K1) cells were obtained from ATCC (CCL61) and maintained at 37°C in a humidified atmosphere (5% CO_2) as a monolayer culture in fenolred-free M505 medium. The latter medium consists of a mixture (1:1) of Dulbecco's Modified Eagle's Medium (DMEM, Gibco 074-200) and Nutrient Medium F12 (Ham's F12, Gibco 074-1700) supplemented with 2.5 mg/ml sodium carbonate (Baker), 55 µg/ml sodium pyruvate (Fluka), 2.3 µg/ml β-mercaptoethanol (Baker), 1.2 µg/ml ethanolamine (Baker), 360 µg/ml L-glutamine (Merck), 0.45 µg/ml sodium selenite (Fluka), 62.5 µg/ml penicillin (Mycopharm), 62.5 µg/ml streptomycin (Serva), and 5% charcoal-treated bovine calf serum (Hyclone).

Recombinant vectors

The ERβ-encoding sequence as presented in SEQ ID No. 1 was amplified by PCR using oligonucleotides 5'-

CTTGGATCCATAGCCCTGCTGTGATGAATTACAG-3' (SEQ ID NO:22 underlined is the translation initiation codon) in combination with 5'-GATGGATCCTCACCTCAGGGCCAGGCGTCACTG-3' (SEQ ID NO:23) (underlined is the translation stopcodon, antisense). The resulting BamH1 fragment (approximately 1450 base pairs) were then cloned in the mammalian cell expression vector pNGV1 (Genbank accession No. X99274).

An expression construct encoding the ERβ reading frame as presented in SEQ ID NO:24 was made by replacing a BamH1-Msc1 fragment (nucleotides 1-81 in SEQ ID No. 1) by a BamH1-Msc1 fragment corresponding to nucleotides 77-316 in SEQ ID No. 24. The latter fragment was made by PCR with SEQ ID NO:26 in combination with SEQ ID NO: 28 using the above mentioned 5' RACE fragment.

The reporter vector was based on the rat oxytocin gene regulatory region (position -363/+16 as a HindIII/ MboI fragment: R.Ivell, and D.Richter, Proc.Natl.Acad.Sci.USA 81, 2006-2010, 1984) linked to the firefly luciferase encoding sequence: the regulatory region of the oxytocin gene was shown to possess functional estrogen hormone response elements *in vitro* for both the rat (R.Adan *et al*, Biochem.Biophys.Res.Comm. 175, 117-122, 1991) and the human (S. Richard and H.Zingg, J.Biol.Chem. 265, 6098-6103, 1990).

Transient transfection

1×10^5 CHO cells were seeded in 6-wells Nunclon tissue culture plates and DNA was introduced by use of lipofectin (Gibco BRL). Hereto, the DNA (1 µg of both receptor and reporter vector in 250 µL Optimem, Gibco BRL) was mixed with an equal volume of lipofectin reagent (7 µL in 250 µL Optimem, Gibco) and allowed to stand at room temperature for 15 min. After washing the cells twice with serum-free medium (M505) new medium (500 µL Optimem, Gibco) was added to the cells followed by the dropwise addition of the DNA-lipofectin mixture. After incubation for a 5 hour period at 37°C cells were washed twice with fenolred-free M505 + 5% charcoal-treated bovine calf serum and incubated overnight at 37°C. After 24 hours hormones were added to the medium (10^{-7} mol/L). Cell extracts were made 48 hours posttransfection by the addition of 200 µL lysisbuffer (0.1 M phosphate buffer pH7.8, 0.2% Triton X-100). After incubation for 5 min at 37°C the cell suspension was centrifuged (Eppendorf centrifuge, 5 min) and 20 µL sample was added to 50 µL luciferase assay reagent (Promega). Light emission was measured in a luminometer (Berthold Biolumat) for 10 sec at 562 nm.

Stable transfection of the novel estrogen receptor.

The expression plasmid encoding full-length ERβ1-530 (see above) was stably transfected in CHO K1 cells as previously described (Theunissen *et al.*, J. Biol. Chem. 268, 9035-9040, 1993). Single cell clones that were obtained this way were screened by transient transfection of the reporter plasmid (rat oxytocin-luciferase) as described above. Selected clones were used for a second stable transfection of the rat oxytocin-luciferase reporter plasmid together with the plasmid pDR2A which contains a hygromycine resitance gene for selection. Single cell clones obtained were tested for a response to 17β-estradiol. Subsequently, a selected single cell clone was used for transactivation studies. Briefly, cells were seeded in 96-wells at (1.6×10^4 cells per well). After 24 hours different concentrations of hormone were diluted in medium and added to the wells. For antagonistic experiments, 2×10^{-10} M. 17β-estradiol was added to each well and different concentrations of antagonists were added. Cells were washed once with PBS after a 24 hour incubation and then lysed by the addition of 40 microliter lysis buffer (see above). Luciferase reagent was added (50 microliter) to each well and light emission was measured using the Topcount (Packard).

Results.

A comparison of the two expression constructs (SEQ ID NO:1 and SEQ ID NO:24) in transient transfections in CHO cells showed identical transactivation in response to a number of agonists and antagonists. CHO cells transiently transfected with ERβ expression vector and a reporter plasmid showed a 3 to 4 fold increase in luciferase activity in response to 17β-estradiol as compared to untreated cells (see Figure 2). A similar transactivation was obtained upon treatment with estriol and estrone. The results indicate not only that the novel ER (ERβ) can bind estrogen hormones but also that the ligand-activated receptor can bind to the estrogen-response elements (EREs) within the rat oxytocin promoter and activate transcription of the luciferase reporter gene. Figure 3 shows that in an independent similar experiment 10^{-9} mol/L 17β-estradiol gave an 18-fold stimulation with ERα and a 7-fold stimulation with ERβ. In addition, the antiestrogen ICI-164384 was shown to be an antagonist for both ERα and ERβ when activated with 17β-estradiol, whereas the antagonist alone had no effect. In this experiment 0.25 µg β-galactosidase vector was co-transfected in order to normalize for differences in transfection efficiency.

Transactivation studies performed on stably transfected ERα and ERβ cell lines gave similar absolute luciferase values. The curves for 17β-estradiol are very similar and show that half-maximal transactivation is reached with lower concentrations of hormone on ERα as compared to ERβ (Figure 5). For Org4094 this is also the case however, the

effect observed is much more pronounced. The curves for raloxifen show that the potency of this antagonist to block transactivation on ERα is greater compared to its potency to block ERβ transactivation

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Akzo nobel n.v.
(B) STREET: Velperweg 76
(C) CITY: Arnhem
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): 6824 BM
(G) TELEPHONE: 0412-666379
(H) TELEFAX: 0412-650592
(I) TELEX: 37503 akpha nl

(ii) TITLE OF INVENTION: Novel estrogen receptor

(iii) NUMBER OF SEQUENCES: 28

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1434 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC 60

ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC 120

CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA 180

GAACACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC 240

GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC 300

GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT 360

AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC 420

GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA 480

ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA 540

AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC 600

CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG 660

GAGGCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC 720

ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC 780

AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC 840

TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG 900

CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT 960

CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAA 1020

CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG 1080

GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG 1140

ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200

CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGTAA CAAGGGCATG 1260

GAACATCTGC TCAACATGAA GTGCAAAAAT GTGGTCCCAG TGTATGACCT GCTGCTGGAG 1320

ATGCTGAATG CCCACGTGCT TCGCGGGTGC AAGTCCTCCA TCACGGGGTC CGAGTGCAGC 1380

CCGGCAGAGG ACAGTAAAAG CAAAGAGGGC TCCCAGAACC CACAGTCTCA GTGA 1434

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1251 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC 60

ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC 120

CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA 180

GAACACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC 240

GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC 300

GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT 360

AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC 420

GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA 480

ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA 540

AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC 600

CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG 660

GAGGCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC 720

ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC 780

AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC 840

TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG 900

CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT 960

CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAA 1020

CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG 1080

GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG 1140

ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200

CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTG A 1251

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp
1   5   10   15

Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His
            20   25   30

Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn
        35   40   45

Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val
    50   55   60

Gly Met
65

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 233 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Leu Val Leu Thr Leu Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser
1   5   10   15

Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser Met Met Met Ser Leu Thr
            20   25   30

Lys Leu Ala Asp Lys Glu Leu Val His Met Ile Ser Trp Ala Lys Lys
        35   40   45

```
Ile Pro Gly Phe Val Glu Leu Ser Leu Phe Asp Gln Val Arg Leu Leu
     50                  55                  60

Glu Ser Cys Trp Met Glu Val Leu Met Met Gly Leu Met Trp Arg Ser
65                  70                  75                  80

Ile Asp His Pro Gly Lys Leu Ile Phe Ala Pro Asp Leu Val Leu Asp
                85                  90                  95

Arg Asp Glu Gly Lys Cys Val Glu Gly Ile Leu Glu Ile Phe Asp Met
            100                 105                 110

Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu Leu Lys Leu Gln His Lys
        115                 120                 125

Glu Tyr Leu Cys Val Lys Ala Met Ile Leu Leu Asn Ser Ser Met Tyr
    130                 135                 140

Pro Leu Val Thr Ala Thr Gln Asp Ala Asp Ser Ser Arg Lys Leu Ala
145                 150                 155                 160

His Leu Leu Asn Ala Val Thr Asp Ala Leu Val Trp Val Ile Ala Lys
                165                 170                 175

Ser Gly Ile Ser Ser Gln Gln Gln Ser Met Arg Leu Ala Asn Leu Leu
            180                 185                 190

Met Leu Leu Ser His Val Arg His Ala Ser Asn Lys Gly Met Glu His
        195                 200                 205

Leu Leu Asn Met Lys Cys Lys Asn Val Val Pro Val Tyr Asp Leu Leu
    210                 215                 220

Leu Glu Met Leu Asn Ala His Val Leu
225                 230
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 477 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Pro
1 5 10 15

Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro Thr Pro Gly His
20 25 30

Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His Leu Tyr Ala Glu
35 40 45

Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu Glu His Thr Leu
50 55 60

Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser Gly Asn Arg Cys
65 70 75 80

Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His Phe Cys
85 90 95

Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser
100 105 110

Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn
115 120 125

Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg
130 135 140

```
Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly
145                 150                 155                 160

Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys Gly Tyr Arg Leu Val
                165                 170                 175

Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala
            180                 185                 190

Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Leu Asp
        195                 200                 205

Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro
    210                 215                 220

Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser
225                 230                 235                 240

Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met
                245                 250                 255

Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
            260                 265                 270

Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
        275                 280                 285

Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
    290                 295                 300

Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile
305                 310                 315                 320

Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
                325                 330                 335

Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
            340                 345                 350
```

Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp
355 360 365

Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu
370 375 380

Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met
385 390 395 400

Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Ser
405 410 415

Asn Lys Gly Met Glu His Leu Leu Asn Met Lys Cys Lys Asn Val Val
420 425 430

Pro Val Tyr Asp Leu Leu Leu Glu Met Leu Asn Ala His Val Leu Arg
435 440 445

Gly Cys Lys Ser Ser Ile Thr Gly Ser Glu Cys Ser Pro Ala Glu Asp
450 455 460

Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln Ser Gln
465 470 475

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 416 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Pro

1 5 10 15

Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro Thr Pro Gly His

20 25 30

Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His Leu Tyr Ala Glu

35 40 45

Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu Glu His Thr Leu

50 55 60

Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser Gly Asn Arg Cys

65 70 75 80

Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His Phe Cys

85 90 95

Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser

100 105 110

Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn

115 120 125

Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg

130 135 140

Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly

145 150 155 160

Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys Gly Tyr Arg Leu Val

165 170 175

Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala

180 185 190

Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Leu Asp

195 200 205

Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro

210 215 220

Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser
225 230 235 240

Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met
245 250 255

Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
260 265 270

Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
275 280 285

Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
290 295 300

Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile
305 310 315 320

Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
325 330 335

Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
340 345 350

Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp
355 360 365

Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu
370 375 380

Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met
385 390 395 400

Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Arg
405 410 415

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGIGAYGARG CWTCIGGITG YCAYTAYGG 29

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAGCCTGGSA YICKYTTIGC CCAIYTIAT 29

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TGTTACGAAG TGGGAATGGT GA 22

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TTGACACCAG ACCAACTGGT AATG 24

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGTGGCGACG ACTCCTGGAG CCCG 24

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GTACACTGAT TTGTAGCTGG AC 22

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CCATGATGAT GTCCCTGACC 20

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TCGCATGCCT GACGTGGGAC 20

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGCSTCCAGC ATCTCCAGSA RCAG 24

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGAAGCTGGC TCACTTGCTG 20

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TCTTGTTCTG GACAGGGATG 20

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GCATGGAACA TCTGCTCAAC 20

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AGCAAGTTCA GCCTGTTAAG T 21

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1257 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC 60

ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC 120

CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA 180

GAACACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC 240

GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC 300

GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT 360

AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC 420

GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA 480

ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA 540

AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC 600

CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG 660

GAGGCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC 720

ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC 780

AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC 840

TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG 900

CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT 960

CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAA 1020

CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG 1080

GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG 1140

ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200

CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTC TGCCTGA 1257

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 418 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Pro
1   5   10   15

Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro Thr Pro Gly His
            20   25   30

Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His Leu Tyr Ala Glu
        35   40   45

Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu Glu His Thr Leu
    50   55   60

Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser Gly Asn Arg Cys
65   70   75   80

Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His Phe Cys
                85   90   95

Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser
            100   105   110

Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn
        115   120   125

Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg
    130   135   140

Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly
145   150   155   160

Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys Gly Tyr Arg Leu Val
                165   170   175

Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala

```
                         180                 185                 190
Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Leu Asp
    195                 200                 205
Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro
    210                 215                 220
Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser
225                 230                 235                 240
Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met
                245                 250                 255
Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
            260                 265                 270
Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
        275                 280                 285
Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
    290                 295                 300
Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile
305                 310                 315                 320
Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
                325                 330                 335
Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
            340                 345                 350
Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp
        355                 360                 365
Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu
    370                 375                 380
Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met
```

385 390 395 400

Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Arg

405 410 415

Ser Ala

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CTTGGATCCA TAGCCCTGCT GTGATGAATT ACAG 34

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GATGGATCCT CACCTCAGGG CCAGGCGTCA CTG 33

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1898 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CACGAATCTT TGAGAACATT ATAATGACCT TTGTGCCTCT TCTTGCAAGG TGTTTTCTCA 60

GCTGTTATCT CAAGACATGG ATATAAAAAA CTCACCATCT AGCCTTAATT CTCCTTCCTC 120

CTACAACTGC AGTCAATCCA TCTTACCCCT GGAGCACGGC TCCATATACA TACCTTCCTC 180

CTATGTAGAC AGCCACCATG AATATCCAGC CATGACATTC TATAGCCCTG CTGTGATGAA 240

TTACAGCATT CCCAGCAATG TCACTAACTT GGAAGGTGGG CCTGGTCGGC AGACCACAAG 300

CCCAAATGTG TTGTGGCCAA CACCTGGGCA CCTTTCTCCT TTAGTGGTCC ATCGCCAGTT 360

ATCACATCTG TATGCGGAAC CTCAAAAGAG TCCCTGGTGT GAAGCAAGAT CGCTAGAACA 420

CACCTTACCT GTAAACAGAG AGACACTGAA AAGGAAGGTT AGTGGGAACC GTTGCGCCAG 480

CCCTGTTACT GGTCCAGGTT CAAAGAGGGA TGCTCACTTC TGCGCTGTCT GCAGCGATTA 540

CGCATCGGGA TATCACTATG GAGTCTGGTC GTGTGAAGGA TGTAAGGCCT TTTTTAAAAG 600

AAGCATTCAA GGACATAATG ATTATATTTG TCCAGCTACA AATCAGTGTA CAATCGATAA 660

AAACCGGCGC AAGAGCTGCC AGGCCTGCCG ACTTCGGAAG TGTTACGAAG TGGGAATGGT 720

GAAGTGTGGC TCCCGGAGAG AGAGATGTGG GTACCGCCTT GTGCGGAGAC AGAGAAGTGC 780

CGACGAGCAG CTGCACTGTG CCGGCAAGGC CAAGAGAAGT GGCGGCCACG CGCCCCGAGT 840

GCGGGAGCTG CTGCTGGACG CCCTGAGCCC CGAGCAGCTA GTGCTCACCC TCCTGGAGGC 900

TGAGCCGCCC CATGTGCTGA TCAGCCGCCC CAGTGCGCCC TTCACCGAGG CCTCCATGAT 960

GATGTCCCTG ACCAAGTTGG CCGACAAGGA GTTGGTACAC ATGATCAGCT GGGCCAAGAA 1020

GATTCCCGGC TTTGTGGAGC TCAGCCTGTT CGACCAAGTG CGGCTCTTGG AGAGCTGTTG 1080

GATGGAGGTG TTAATGATGG GGCTGATGTG GCGCTCAATT GACCACCCCG GCAAGCTCAT 1140

CTTTGCTCCA GATCTTGTTC TGGACAGGGA TGAGGGGAAA TGCGTAGAAG GAATTCTGGA 1200

AATCTTTGAC ATGCTCCTGG CAACTACTTC AAGGTTTCGA GAGTTAAAAC TCCAACACAA 1260

AGAATATCTC TGTGTCAAGG CCATGATCCT GCTCAATTCC AGTATGTACC CTCTGGTCAC 1320

AGCGACCCAG GATGCTGACA GCAGCCGGAA GCTGGCTCAC TTGCTGAACG CCGTGACCGA 1380

TGCTTTGGTT TGGGTGATTG CCAAGAGCGG CATCTCCTCC CAGCAGCAAT CCATGCGCCT 1440

GGCTAACCTC CTGATGCTCC TGTCCCACGT CAGGCATGCG AGTAACAAGG GCATGGAACA 1500

TCTGCTCAAC ATGAAGTGCA AAAATGTGGT CCCAGTGTAT GACCTGCTGC TGGAGATGCT 1560

GAATGCCCAC GTGCTTCGCG GGTGCAAGTC CTCCATCACG GGGTCCGAGT GCAGCCCGGC 1620

AGAGGACAGT AAAAGCAAAG AGGGCTCCCA GAACCCACAG TCTCAGTGAC GCCTGGCCCT 1680

GAGGTGAACT GGCCCACAGA GGTCACAAGC TGAAGCGTGA ACTCCAGTGT GTCAGGAGCC 1740

TGGGCTTCAT CTTTCTGCTG TGTGGTCCCT CATTTGGTGA TGGCAGGCTT GGTCATGTAC 1800

CATCCTTCCC TCCACCTTCC CAACTCTCAG GAGTCGGTGT GAGGAAGCCA TAGTTTCCCT 1860

TGTTAGCAGA GGGACATTTG AATCGAGCGT TTCCACAC 1898

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 530 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Met Asp Ile Lys Asn Ser Pro Ser Ser Leu Asn Ser Pro Ser Ser Tyr
1 5 10 15

Asn Cys Ser Gln Ser Ile Leu Pro Leu Glu His Gly Ser Ile Tyr Ile
20 25 30

Pro Ser Ser Tyr Val Asp Ser His His Glu Tyr Pro Ala Met Thr Phe
35 40 45

Tyr Ser Pro Ala Val Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn
50 55 60

Leu Glu Gly Gly Pro Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp
65 70 75 80

Pro Thr Pro Gly His Leu Ser Pro Leu Val Val His Arg Gln Leu Ser
85 90 95

His Leu Tyr Ala Glu Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser
100 105 110

Leu Glu His Thr Leu Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val

115 120 125

Ser Gly Asn Arg Cys Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg
130 135 140

Asp Ala His Phe Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His
145 150 155 160

Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser
165 170 175

Ile Gln Gly His Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr
180 185 190

Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys
195 200 205

Cys Tyr Glu Val Gly Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys
210 215 220

Gly Tyr Arg Leu Val Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His
225 230 235 240

Cys Ala Gly Lys Ala Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg
245 250 255

Glu Leu Leu Leu Asp Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu
260 265 270

Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro
275 280 285

Phe Thr Glu Ala Ser Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys
290 295 300

Glu Leu Val His Met Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val
305 310 315 320

Glu Leu Ser Leu Phe Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met

325 330 335

Glu Val Leu Met Met Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly
340 345 350

Lys Leu Ile Phe Ala Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys
355 360 365

Cys Val Glu Gly Ile Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr
370 375 380

Ser Arg Phe Arg Glu Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val
385 390 395 400

Lys Ala Met Ile Leu Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala
405 410 415

Thr Gln Asp Ala Asp Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala
420 425 430

Val Thr Asp Ala Leu Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser
435 440 445

Gln Gln Gln Ser Met Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His
450 455 460

Val Arg His Ala Ser Asn Lys Gly Met Glu His Leu Leu Asn Met Lys
465 470 475 480

Cys Lys Asn Val Val Pro Val Tyr Asp Leu Leu Leu Glu Met Leu Asn
485 490 495

Ala His Val Leu Arg Gly Cys Lys Ser Ser Ile Thr Gly Ser Glu Cys
500 505 510

Ser Pro Ala Glu Asp Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln
515 520 525

Ser Gln

530

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GTGCGGATCC TCTCAAGACA TGGATATAAA 30

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

AGTAACAGGG CTGGCGCAAC GGTTC 25

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

ACTGGCGATG GACCACTAAA GG 22

Claims

1. Isolated DNA encoding a protein having an N-terminal domain, a DNA-binding domain and a ligand-binding domain, wherein the amino acid sequence of said DNA-binding domain of said protein exhibits at least 80% homology with the amino acid sequence shown in SEQ ID NO:3, and the amino acid sequence of said ligand-binding domain of said protein exhibits at least 70% homology with the amino acid sequence shown in SEQ ID NO:4.

2. Isolated DNA according to claims 1, characterized in that the amino acid sequence of said DNA-binding domain of said protein exhibits at least 90%, preferably 95%, more preferably 98%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO:3.

3. Isolated DNA according to claims 1 or 2, characterized in that the amino acid sequence of said ligand-binding domain of said protein exhibits at least 75%, preferably 80%, more preferably 90%, most preferably 100% homology with the amino acid sequence shown in SEQ ID NO:4.

4. Isolated DNA according to claims 1 to 3, said DNA encoding a protein comprising the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.

5. Isolated DNA according to claims 1 to 4, characterized in that said DNA comprises the nucleic acid sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:20 or SEQ ID NO:24.

6. A recombinant expression vector comprising the DNA according to any of the claims 1 to 5.

7. A cell transfected with DNA according to claims 1 to 5 or an expression vector according to claim 6.

8. A cell according to claim 7 which is a stable transfected cell line which expresses the steroid receptor protein according to any of the claims 9 to 11.

9. Protein encoded by DNA according to claims 1 to 5 or an expression vector according to claim 6.

10. Protein according to claim 9, said protein comprising the amino acid sequence of SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:21 or SEQ ID NO:25.

11. Chimeric protein having an N-terminal domain, a DNA-binding domain, and a ligand-binding domain,

characterized in that at least one of said domains of said chimeric protein originates from a protein according to claims 9 or 10, and at least one of the other domains of said chimeric protein originates from another receptor protein from the nuclear receptor superfamily, provided that the DNA-binding domain and the ligand-binding domain of said chimeric protein originates from different proteins.

12. DNA encoding a protein according to claim 11.

13. Use of a DNA according to claims 1 to 5 or 12, an expression vector according to claim 6, a cell according to claim 7 or 8 or a protein according to claim 9 to 11 in a screening assay for identification of new drugs.

14. A method for identifying functional ligands for the protein according to claims 9 to 11, said method comprising the steps of

    a) introducing into a suitable host cell 1) DNA according to claims 1 to 5 or 12, and 2) a suitable reporter gene functionally linked to an operative hormone response element, said HRE being able to be activated by the DNA-binding domain of the protein encoded by said DNA;
    b) bringing the host cell from step a) into contact with potential ligands which will possibly bind to the ligand-binding domain of the protein encoded by said DNA from step a);
    c) monitoring the expression of the protein encoded by said reporter gene of step a).

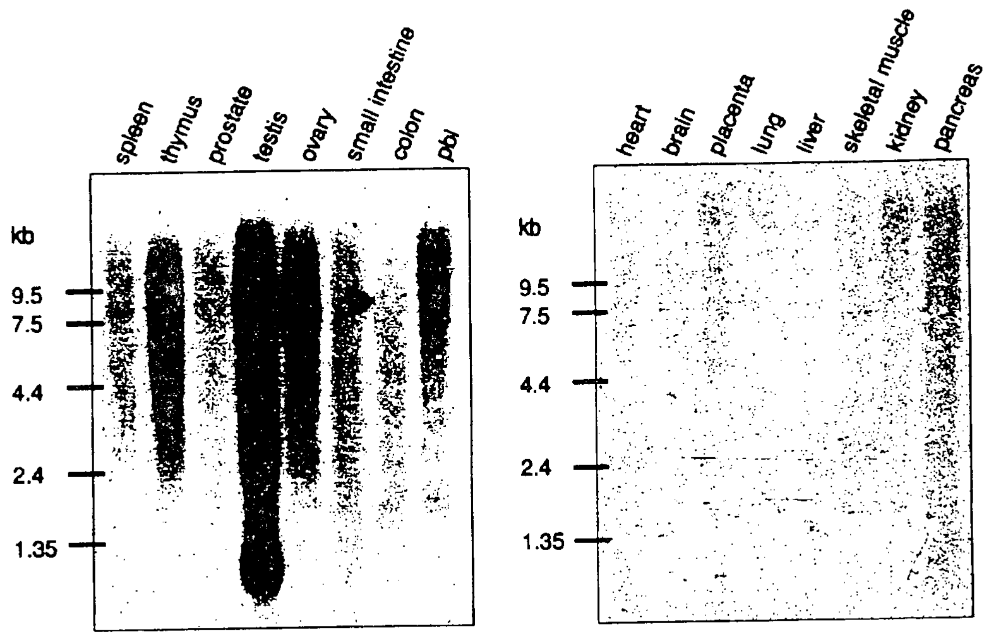
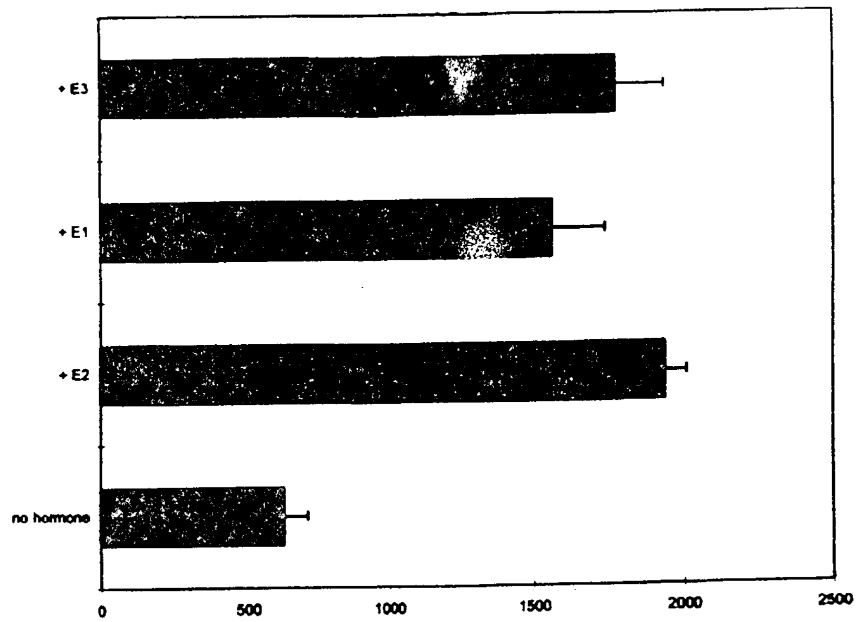

spleen
thymus
prostate
testis
ovary
small intestine
colon
pbl
kb
9.5
7.5
4.4
2.4
1.35
heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
kb
9.5
7.5
4.4
2.4
1.35

Figure 1



luciferase units

Fig. 2

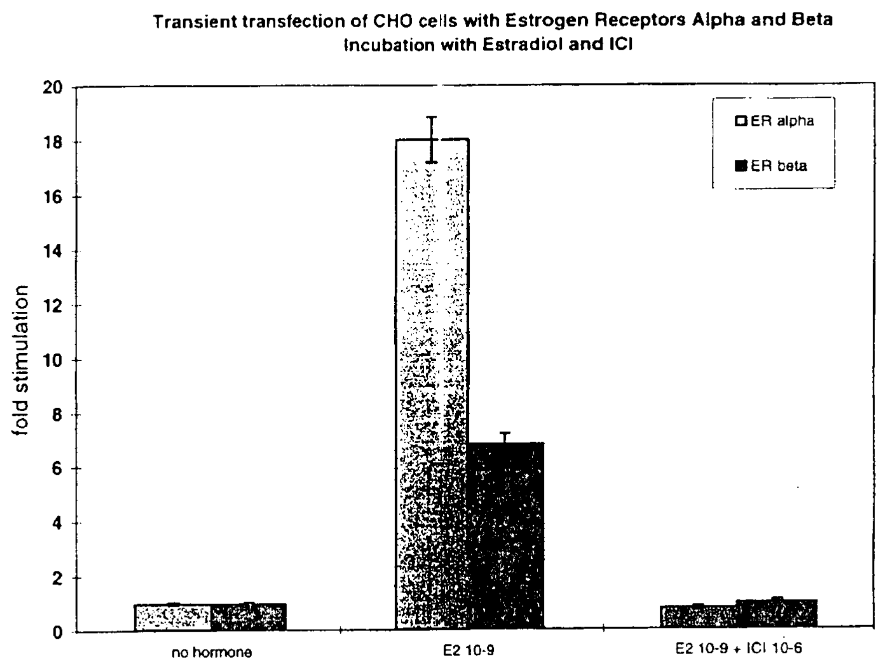
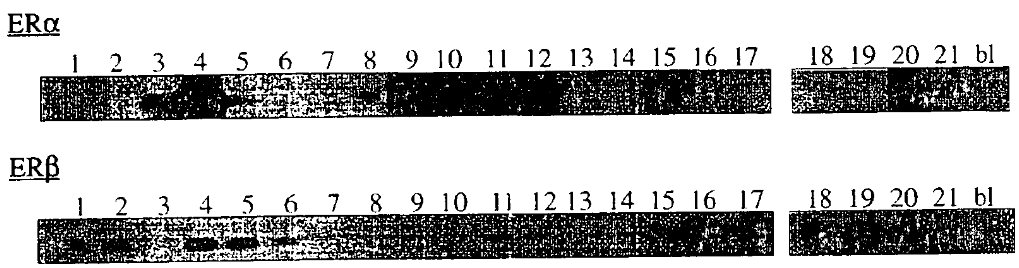

Transient transfection of CHO cells with Estrogen Receptors Alpha and Beta
Incubation with Estradiol and ICI
ER alpha
ER beta
fold stimulation
20
18
16
14
12
10
8
6
4
2
0
no hormone
E2 10-9
E2 10-9 + ICI 10-6

Figure 3

ERα and ERβ RT PCR on tissue-representative cell lines



1. Ishikawa
2. HEC-1A
3. RL95-2
4. ECC-1
5. SaOS-2
6. HOS
7. U2-OS
8. MG-63
9. MCF-7
10. T47-D
11. Hs-760T
12. SW-954
13. Hep-G2
14. CaCo
15. HISM
16. HUV-EC-C
17. BAEC-1
18. A10
19. A7R5
20. CavaSMC
21. RASMC

bl. blank

Figure 4

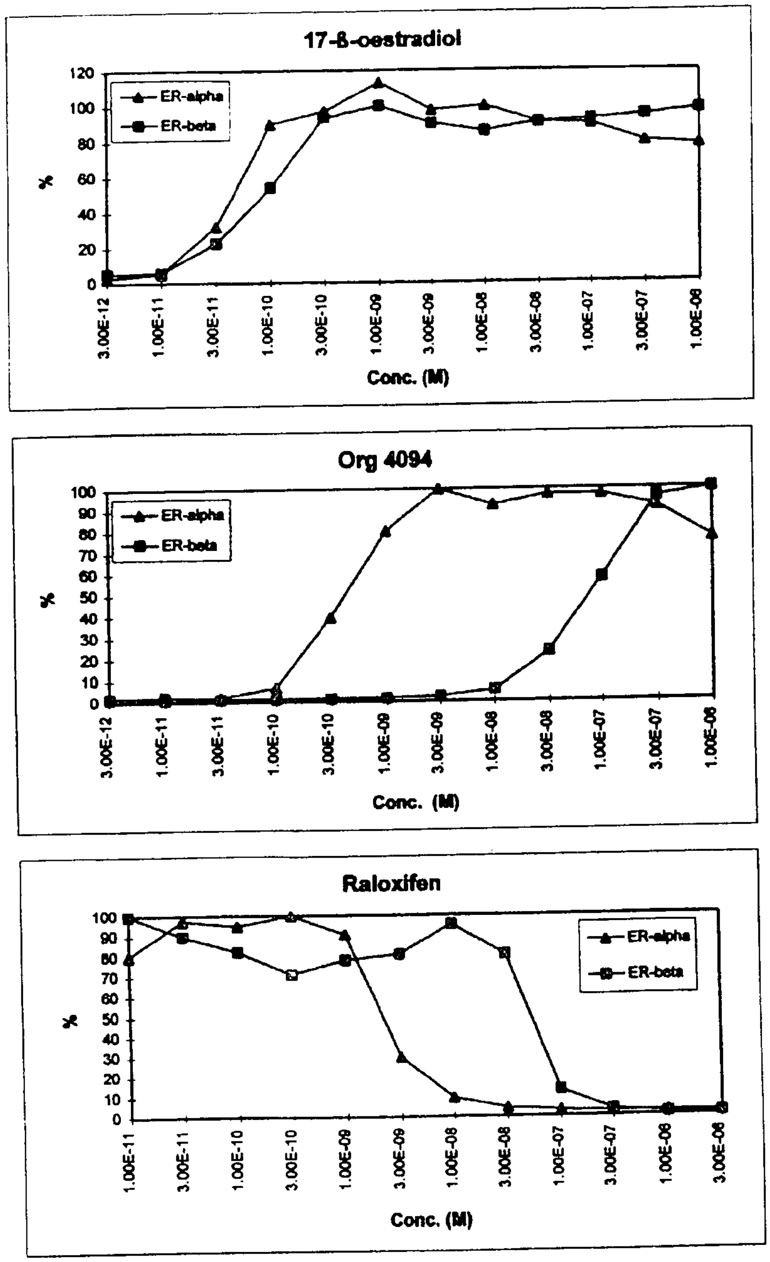

17-ß-oestradiol
ER-alpha
ER-beta
%
120
100
80
60
40
20
0
3.00E-12
1.00E-11
3.00E-11
1.00E-10
3.00E-10
1.00E-09
3.00E-09
1.00E-08
3.00E-08
1.00E-07
3.00E-07
1.00E-06
Conc. (M)
Org 4094
ER-alpha
ER-beta
%
100
90
80
70
60
50
40
30
20
10
0
3.00E-12
1.00E-11
3.00E-11
1.00E-10
3.00E-10
1.00E-09
3.00E-09
1.00E-08
3.00E-08
1.00E-07
3.00E-07
1.00E-06
Conc. (M)
Raloxifen
ER-alpha
ER-beta
%
100
90
80
70
60
50
40
30
20
10
0
1.00E-11
3.00E-11
1.00E-10
3.00E-10
1.00E-09
3.00E-09
1.00E-08
3.00E-08
1.00E-07
3.00E-07
1.00E-06
3.00E-06
Conc. (M)

Figure 5

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 798 378 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
29.12.1997 Bulletin 1997/52

(43) Date of publication A2:
01.10.1997 Bulletin 1997/40

(21) Application number: 97200903.9

(22) Date of filing: 25.03.1997

(51) Int Cl.6: C12N 15/12, C12N 15/62, C12N 15/85, C07K 14/72, C12N 1/21, C12N 5/16, C12Q 1/00, C12Q 1/68

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: 22.11.1996 EP 96203284
26.03.1996 EP 96200820

(71) Applicant: Akzo Nobel N.V.
6824 BM Arnhem (NL)

(72) Inventors:
- Mosselman, Sietse
  5346 VM OSS (NL)
- Dijkema, Rein
  5345 ML Oss (NL)

(74) Representative:
Ogilvie-Emanuelson, Claudia Maria et al
Patent Department Pharma
N.V. Organon
P.O. Box 20
5340 BH Oss (NL)

(54) **Estrogen receptor**

(57) The present invention relates to isolated DNA encoding novel estrogen receptors, the proteins encoded by said DNA, chimeric receptors comprising parts of said novel receptors and uses thereof.

EP 0 798 378 A3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 20 0903

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y,D	GREEN, G.L.: "Sequence and expression of human estrogen receptor complementary DNA." SCIENCE, vol. 231, 13 March 1986, pages 1150-1154, XP000611679 * the whole document, especially Figure 1 *	1-14	C12N15/12 C12N15/62 C12N15/85 C07K14/72 C12N1/21 C12N5/16 C12Q1/00 C12Q1/68
Y	CHANG, C. ET AL.: "Human and rat TR4 orphan receptors specify a subclass of the steroid receptor superfamily." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE, vol. 91, June 1994, pages 6040-6044, XP000606867 * the whole document, especially 'Materials and Methods' (PCR cloning of ..) and Figure 1 *	1-14	
A,D	EVANS, R.M.: "The steroid and thyroid hormone receptor superfamily." SCIENCE, vol. 240, 13 May 1988, pages 889-895, XP002019515 * the whole document *	1-14	TECHNICAL FIELDS SEARCHED (Int.Cl.6) C12N C07K C12Q
A	EP 0 371 820 A (SALK INST FOR BIOLOGICAL STUDI) 6 June 1990 * page 4, lines 20-27 and claim 19 *	11-14	
E	WO 97 09348 A (KAROBIO AB ;KUIPER GEORGE G J M (SE); ENMARK EVA (SE); GUSTAFSSON) * the whole document *	1-14	

-/--

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	23 October 1997	Mandl, B

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 20 0903

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y,D	GREEN, G.L.: "Sequence and expression of human estrogen receptor complementary DNA." SCIENCE, vol. 231, 13 March 1986, pages 1150-1154, XP000611679 * the whole document, especially Figure 1 *	1-14	C12N15/12 C12N15/62 C12N15/85 C07K14/72 C12N1/21 C12N5/16 C12Q1/00 C12Q1/68
Y	CHANG, C. ET AL.: "Human and rat TR4 orphan receptors specify a subclass of the steroid receptor superfamily." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE, vol. 91, June 1994, pages 6040-6044, XP000606867 * the whole document, especially 'Materials and Methods' (PCR cloning of ..) and Figure 1 *	1-14	
A,D	EVANS, R.M.: "The steroid and thyroid hormone receptor superfamily." SCIENCE, vol. 240, 13 May 1988, pages 889-895, XP002019515 * the whole document *	1-14	TECHNICAL FIELDS SEARCHED (Int.Cl.6) C12N C07K C12Q
A	EP 0 371 820 A (SALK INST FOR BIOLOGICAL STUDI) 6 June 1990 * page 4, lines 20-27 and claim 19 *	11-14	
E	WO 97 09348 A (KAROBIO AB ;KUIPER GEORGE G J M (SE); ENMARK EVA (SE); GUSTAFSSON) * the whole document *	1-14	

-/--

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	23 October 1997	Mandl, B

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 20 0903

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
P,X	MOSSELMAN S. ET AL.: "ER-beta: Identification and characterization of a novel human estrogen receptor." FEBS LETTERS, vol. 392, no. 1, 19 August 1996, pages 49-53, XP002044410 * the whole document * -----	1-10	

TECHNICAL FIELDS SEARCHED (Int.Cl.6)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	23 October 1997	Mandl, B

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)